Heterogeneous Firms, ‘Profit Shifting’ FDI and International Tax Competition*

Sebastian Krautheim
European University Institute and Paris School of Economics

Tim Schmidt-Eisenlohr
European University Institute

December 12, 2008
PRELIMINARY DRAFT - NOT FOR CIRCULATION

Abstract

We propose a stylized model of international tax competition between a large country and a tax haven. In the large country, firms in a monopolistically competitive industry generate positive profits which can be taxed by the government. Firms have heterogeneous productivity levels and can choose to undertake ‘profit shifting’ FDI in order to benefit from lower tax rates abroad. The methodological contribution of our paper is to provide a fully solvable model of international tax competition with heterogeneous firms and monopolistic competition. In equilibrium the most productive firms are the most likely to shift profits abroad. Whether these firms account for a large fraction of aggregate profits depends on firm heterogeneity and monopolistic market power. The analysis reveals that economies with a low degree of firm heterogeneity (relatively few very productive firms) and a high degree of monopolistic market power (low substitutability between goods) are ‘shielded’ from international tax competition. In equilibrium they face lower outflows of the tax base and can set higher tax rates.

JEL: F23, H25, H87

Keywords: heterogeneous firms, monopolistic competition, tax competition, tax havens

* We are grateful to Giancarlo Corsetti, Morten Ravn and Omar Licandro for constant advice. We would also like to thank guest researchers at the Deutsche Bundesbank for fruitful discussions and comments. Any remaining errors are ours.
1 Introduction

International financial integration enables multinational companies to use international tax planning strategies to reduce their tax payments. A direct consequence is that when setting their tax policies governments are competing for a potentially mobile tax base.

Very few contributions in the economics literature have considered the role industry structure (i.e., firm heterogeneity) plays for the analysis of international tax competition.\(^1\) Recent empirical evidence shows, however, that firm heterogeneity does matter for the extent of international tax optimization. A recent example is Desai, Foley, and Hines (2006). Using affiliate-level data of American firms, they show that firm size, internationalization and R&D intensity increase the probability of using tax haven operations.

This paper proposes a fully solvable model of international tax competition with heterogeneous firms and monopolistic competition. The model allows to analyze the impact of firm heterogeneity and monopolistic market power on the main policy tradeoffs in international tax competition. The analysis reveals that firm heterogeneity (together with the degree of monopolistic market power) determines the distribution of profits (the tax base) across firms. This distribution is key for optimal government policies as high productivity firms are more likely to react to tax differentials than low productivity firms. The analysis shows that the effects of international tax competition are strongest in economies with a low degree of firm heterogeneity (relatively many productive firms) and high substitutability across goods (low monopolistic market power).

The model features a large country, in which firms are active in a monopolistically competitive industry and earn positive profits. These profits can be taxed by the government. In order to avoid taxation at home, firms can become multinationals by opening an affiliate in a foreign jurisdiction (the ‘tax haven’). This tax haven is small in the sense that it does not have an industry of its own producing differentiated products. Using methods of profit shifting, multinationals can transfer all profits to their affiliates and thus effectively pay taxes in the tax haven only.\(^2\)

A crucial feature of the model is that, as in the data, firms have heterogeneous productivity levels, equilibrium sizes and profits. The introduction of heterogeneity allows us to analyze the empirically more relevant profit taxes rather than a per unit capital taxes. STIMMT DAS?

In the unique Nash-equilibrium of the model, the tax haven sets a tax rate below the one set by

---

\(^1\)Some notable exceptions will be discussed below.

\(^2\)In order to keep the analysis focused, we only allow for this ‘profit shifting’ aspect of FDI. We only consider the case where firms can shift their total profits abroad. Introducing partial profit shifting would neither affect the mechanisms in the model nor the basic results.
the large country.\footnote{This unique equilibrium exists under a condition on the parameters of the model that will be discussed in the main text.} This gives firms an incentive to do ‘profit shifting’ FDI. While the fixed cost of opening an affiliate in the tax haven is the same for all firms, the gains from profit shifting depend on the level of profits a firm is making. In line with the empirical evidence on international tax planning (see e.g. Desai, Foley, and Hines (2006)) in equilibrium the most productive (and thus largest and most profitable) firms become multinationals while less productive firms continue to pay taxes at home.

The higher the tax rate of the large country (relative to the tax rate of the tax haven), the more firms will create an affiliate in the tax haven to pay taxes abroad. An increase in the tax rate in the large country thus has two opposing effects on tax revenues at home. On the one hand it increases revenues per unit of profits taxed (intensive margin), on the other hand it decreases the size of the tax base itself by making tax evasion more attractive (extensive margin).

A key determinant of the optimal tax policies is the distribution of profits across firms with different productivity levels. The most productive firms earn the highest individual profits and are thus the first to do ‘profit shifting’ FDI. For the reaction of the tax base (aggregate profits) to tax differences it is crucial for how much of aggregate profits these firms account. This is determined by industry- and market structure. When the economy features a high degree of firm heterogeneity, there is a higher number of relatively productive firms. Since these firms are (relatively) numerous they also account for a larger share of aggregate profits. In addition, these high productivity firms have high individual profits when the degree of substitutability between goods is high (low monopolistic market power). When this is the case these firms benefit more (in terms of sales and profits) from their productivity advantage over low productivity firms.

By how much the tax haven undercuts the large country is determined by the elasticity of the tax base it is facing. This elasticity depends on firm heterogeneity, monopolistic market power and the levels of the tax rates. For a given tax rate of the large country this elasticity is increasing in the tax haven’s tax rate. The tax haven maximizes its total tax revenues. Thus (starting from the large country’s tax rate where the elasticity would be infinity) it lowers its tax rate until the intensive and extensive margin effect offset each other and the elasticity of the tax base equals unity. When the most productive firms account for a large fraction of aggregate profits (high heterogeneity and low market power) only a low tax difference is necessary to reach an elasticity of unity. Thus, when firm heterogeneity is high and market power low, the tax haven will not strongly undercut the large country, because already a small tax difference induces a high inflow
It is an important result of this paper that the industry structure in an economy crucially determines how strongly it is affected by international tax competition. Our analysis shows that the equilibrium outflows of the tax base to the tax haven are larger the higher the degree of heterogeneity in the economy. Put differently, a lower degree of firm heterogeneity ‘shields’ countries from international tax competition and allows them to set higher tax rates.

This paper relates to two different strands of the literature on international tax competition. One is the Public Finance literature on international tax competition which builds on the seminal contributions of Zodrow and Mieszkowski (1986) and Wilson (1986) (see e.g. Wilson (1999) or Fuest, Huber, and Mintz (2005) for surveys). This literature mainly uses models with identical firms under perfect competition. Usually there are two factors of production: labor and capital. The former is in general assumed to be immobile and the latter to some degree mobile between countries. A benevolent government sets a capital tax rate to maximize welfare of its citizens. It then faces a trade-off between budget for public good provision and the capital outflow implied by high tax rates. In equilibrium governments set inefficiently low tax rates.

We consider a tax game between the large country and the tax haven which follows this literature to some extent. In particular we allow for one single tax instrument (a corporate tax) only. We do, however, not consider per unit capital taxes, as standard in this literature, but analyze a proportional profit tax, which we believe to be the relevant instrument to study.

Building on this literature, Burbidge, Cuff, and Leach (2006) introduce a type of firm heterogeneity into a model with perfect competition, immobile labor and mobile capital. They model firm heterogeneity as an idiosyncratic, exogenous comparative advantage in one of the locations. So firms are heterogeneous in the sense that they are more productive in one country or the other. This is reflected by a country-firm specific TFP term in a decreasing returns to scale production function. Their analysis shows that these location-specific productivities can lead firms owned by residents of the home country to locate in the foreign market. The type of heterogeneity they use is fundamentally different from the ‘Melitz (2003)’-type of heterogeneity where (independent of the location of their production) firms differ in productivity and thus sales and profits.

Our model is also related to the literature on tax competition in a ‘New Economic Geography’ (NEG) context (see e.g. Baldwin and Krugman (2004), Boreck and Pflueger (2006) and Ottaviano and van Ypersele (2005)). The analysis of NEG models is usually focused on the effect of taxation and tax competition on the (re)allocation of monopolistically competitive firms be-
tween countries of different size. Due to the well known ‘home-market effect’ (Krugman (1980)), the larger country attracts more firms and can thus afford to set higher taxes on capital than the smaller country. The larger country thereby taxes capital owned by agents in the foreign market.\footnote{Sato and Thisse (2007) show that in a setting with imperfect matching between firms and workers, this home market effect can be reversed.}

At first sight it might seem to be a straight forward generalization of NEG models of tax competition to introduce firm heterogeneity as in Melitz (2003). To the best of our knowledge, however, there have not been many attempts to do so in the literature. Baldwin and Okubo (2008) outline how a basic setup of an NEG model with tax competition and heterogeneous firms could look like. They do, however, not derive the equilibrium of their model. Instead, they focus their analysis on the investigation of a trade-off between two different tax instruments for an exogenous (possibly off-equilibrium) tax difference.

Davies and Eckel (2007) also propose an NEG-type model of tax competition with heterogeneous firms. In order to achieve tractability they make a strong assumption on the ownership structure of firms (and thus on the effect of profits - and thereby taxation - on demand). They assume that firms belong to individual entrepreneurs which do not have an individual demand of their own. Once entrepreneurs change the country (to maximize profits of their firms), they ‘join’ the representative consumer and then share their firms’ profits with all citizens of the country. So firms can be seen as independent units that can choose freely to move to the country in which they have the highest profits. But once they move, all the profits of the firm go to the representative consumer in the host country.

We view it as a methodological contribution of our paper to provide a (to our knowledge the first) tractable general equilibrium model of international tax competition with heterogeneous firms and monopolistic competition. To achieve tractability, we abstract from several issues that are crucial in the analysis of the NEG literature, but not in our context: we consider a case of ‘very’ asymmetric countries (large country and tax haven). We are not analyzing firm creation or endogenous (re)location of firms but focus on existing firms. We abstract from international trade and focus on ‘profit shifting’ FDI. Furthermore, we assume quasi-linear preferences (which are linear in a homogeneous good and the public good). In this setup both the total expenditure on differentiated goods and the aggregate price level are independent of the tax rates. While these are strong simplifications, we believe them to be appropriate for the purpose of the paper. They allow us to solve the model in closed form and to investigate the role of firm heterogeneity
and monopolistic market power in international tax competition further than has been done in the literature so far.

The remainder of the paper is structured as follows. Section two presents the case of a large country in financial autarky. In section three the tax haven and ‘profit shifting’ FDI are introduced. Section four analyzes the equilibrium. Section five discusses the main results of the paper. In section six we discuss the assumptions necessary to achieve tractability. Section seven concludes.

2 Financial Autarky

We first outline the structure of the ‘Home’ economy in financial autarky. The economy in the large country is endowed with $L$ workers. There are two sectors, one producing varieties of a differentiated good and one producing a homogeneous good with constant returns to scale. The homogeneous good is freely traded and is used as the numeraire with its price normalized to one. As is standard in such a setting (see, for example, Melitz and Ottaviano (2008)) only those equilibria are considered where homogeneous good is produced implying that wages are equalized across countries and can be normalized to unity. Labor is supplied inelastically and is the only input in production. There is a fixed and exogenous measure of firms in the large countries $\Omega$, which are owned by consumers in the large country.

Preferences: The workers are all identical and share the same quasi-linear preferences over consumption of the two goods and a (private) good provided by the government:

$$U = \alpha \ln Q + \beta G + q_0 \quad \text{with} \quad Q = \left( \int_{\Omega} q(\omega)^{\frac{\alpha}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

Where $q(\omega)$ is the quantity consumed of variety $\omega$. The elasticity of substitution between varieties is given by $\sigma > 1$ and $Q$ thus represents consumption of a preference weighted basket of differentiated goods. $G$ is the quantity of a public good provided the government. The consumption of the numeraire good is given by $q_0$. $\alpha$ and $\beta$ are parameters with $0 < \alpha < 1 < \beta$.

Taking prices and available varieties as given, the demand for one particular variety takes the

---

5 Alternatively one could introduce an initial endowment of the homogeneous good that is large enough to assure that both goods are consumed.
usual form:

\[ q(\omega) = \frac{p(\omega)^{1-\sigma}}{P^{1-\sigma}} \alpha \]  

Where \( p(\omega) \) is the price of variety \( \omega \), the aggregate price index of the differentiated goods sector in given by \( P = \left( \int_0^1 p(a)^{1-\sigma} dG(a) \right)^{1-\sigma} \) and \( \alpha = P Q \) is the overall expenditure on goods in this sector.

**The government:** The only tax instrument of the government is a profit tax on the profits of firms in the home country.\(^6\) Tax income can be used to provide government services \( G \) to the consumers. The government can transform one unit of the numeraire good into one unit of the government services. The government is assumed to maximize welfare of its own citizens.

**Firms:** In the homogeneous good sector firms produce with a constant returns to scale technology and earn zero profits.

Similar to Chaney (2008), there is a fixed and exogenous measure of firms in the differentiated good sector that is without loss of generality normalized to one. Each firm produces a different variety. Firms differ in their levels of marginal cost, which is constant for each firm. We assume that these marginal cost levels reach over an interval of \([0, a_{\text{max}}]\) and are following a Pareto distribution with the distribution function given by

\[ G(a) = \left( \frac{a}{a_{\text{max}}} \right)^\gamma \]

with \( a_{\text{max}} \) defined as the highest marginal cost (i.e. the lowest marginal productivity). Without loss of generality we normalize \( a_{\text{max}} = 1 \). The shape parameter \( \gamma \) of the Pareto distribution determines the degree of heterogeneity.\(^7\) As standard in the literature, we assume \( \gamma > (\sigma - 1) \) in order to assure a finite mean of the productivity distribution.\(^8\) There is no fixed cost of production for firms, so that in equilibrium all firms will produce.\(^9\)

---

\(^6\)This is a stylized way of looking at corporate taxation based on the ‘source principal’. By allowing for one tax instrument only, we place our model in the tradition of the large literature on tax competition surveyed by Wilson (1999). Following Hamada (1966), some authors add a lump sum tax on labor, which usually leads to negative tax rates on capital (see for example Burbidge, DePater, Myers, and Sengupta (1997) or Ottaviano and van Ypersele (2005)).

\(^7\)When \( \gamma \) is high, a large mass of firms will have high cost levels and only a small fraction will have low cost levels. In this case we will speak of a low degree of heterogeneity, because ‘on average’ firms are very similar.

\(^8\)Here, productivity is defined as \( 1/a \).

\(^9\)The implications and advantages of this assumption will be discussed below.
Firms in the differentiated good sector charge a constant mark-up over marginal cost:

\[ p(a) = \frac{\sigma}{\sigma - 1} \alpha. \]  

(3)

The level of the mark-up depends on the elasticity of substitution between varieties. When \( \sigma \) is high, firms have a low degree of monopolistic market power and can only afford to charge a low mark-up.

A firm’s gross profits are given by \( \pi(a) = r(a)/\sigma \) which implies

\[ \pi(a) = a^{1-\sigma} T_1 \quad \text{with} \quad T_1 = \frac{(\frac{\sigma - 1}{\sigma})^{\sigma - 1}}{\alpha P^{\sigma - 1}/\sigma} \]  

(4)

\( T_1 \) is a constant that only depends on parameters of the model and the price index, which is also constant.\(^{10}\)

Net profits are just given by \( \pi(a)_{\text{net}} = (1 - t) \pi(a) \), where \( t \in [0, 1] \) is a tax rate set by the government and taken to be exogenous by the firm. Firm choices that maximize gross profits also maximize net profits. The tax is thus not distorting the optimal behavior of the firms.\(^{11}\)

**The tax base - aggregate profits:** In financial autarky all firms pay taxes at home. The tax base is thus given by aggregate profits of domestic firms \( \Pi_H = \int_0^1 \pi(a) \, dG(a) \). Evaluating the integral using (4) leads to

\[ \Pi_H^A = \frac{\epsilon + 1}{\epsilon} T_1 = \frac{\alpha}{\sigma} \]  

(5)

or, expressed differently for later reference, \( \Pi_H^A = T_2^{-\epsilon} T_1^{-\epsilon + 1} \). For notational convenience we have defined

\[ \epsilon \equiv \frac{\gamma}{\sigma - 1} - 1 \quad \text{and} \quad T_2 \equiv \frac{\epsilon + 1}{\epsilon} T_1^{-\epsilon + 1} \]

It is of interest to note that \( \epsilon \) represents two of the crucial parameters of the model: the degree of firm heterogeneity and the elasticity of substitution between varieties. Recall that above, we have assumed that \( \gamma > (\sigma - 1) > 1 \) which implies \( \epsilon > 0 \).

\(^{10}\)The price index is defined as \( P = \left( \int_0^1 p(a)^{1-\sigma} dG(a) \right)^{1/1-\sigma} \). Evaluating the integral using (3) leads to \( P = \frac{\sigma}{\sigma - 1} \left( \frac{\sigma - 1}{\sigma} \right)^{\sigma - 1} \) which is a constant. Using this expression we find \( T_1 = \frac{\alpha}{\sigma} \gamma (\sigma - 1) \).

\(^{11}\)If we allowed for firm entry and exit this statement would still be true for active firms, but the tax rate would affect the set of active firms.
Optimal tax rate in autarky: Households have income from labor and receive the net profits of firms in their country. In autarky the aggregate income $I^A$ of consumers is thus

$$I^A = L + (1 - t_H^A) \Pi_H^A.$$  

Welfare in financial autarky is then given by:

$$U^A = \bar{U} + (1 - t_H^A) \Pi_H^A + \beta \, t_H^A \Pi_H^A.$$  

(6)

Where $\bar{U} \equiv \alpha \ln \left( \frac{a}{p} \right) - \alpha + L$ collects terms that are unaffected by the taxation decision. The first term in $\bar{U}$ reflects utility of consuming the basket of differentiated products (which is constant due to our assumptions on preferences and endowments), $\alpha$ reflects the cost of consuming the basket of differentiated products in terms of forgone consumption of the homogeneous good. $L$ is labor income which maps one to one into utility from consuming the homogeneous good. The last term in (6) represent the part of aggregate firm profits that is paid to the government and transformed into the government service. The profits retained by consumers are used for consumption of the numeraire good.

The following proposition summarizes the main results of this section.

**Proposition 1** In financial autarky the welfare maximizing tax rate of the large country is given by $t_H^A = 1$.

**Proof:** This follows directly from the fact that $\bar{U}$ is constant, $\Pi_H^A > 0$ and $\beta > 1$. q.e.d.

While this result is of course very stylized, it provides a good benchmark for the analysis of the tax haven below.

3 Introducing ‘profit shifting’ FDI

3.1 Financial Integration

We now consider the case where the large country described above is financially integrated with a small independent jurisdiction: the tax haven. Financial integration in our context means that firms in the large country have the possibility to become a multinational by opening an affiliate in the tax haven and shift their profits from the headquarters to the affiliate. These profits are then taxed according to the tax rate in the tax haven, but not at home, where the firm declares zero profits. Opening an affiliate in the tax haven requires paying a fixed cost $f_t$. 
As before the government in the large country maximizes welfare of its citizens. For simplicity, we assume that the tax haven has no tax base of its own and just maximizes tax revenues from taxing foreign firms’ profits which are transferred to the tax haven via FDI. Just as in financial autarky, the only variable the governments can set are the profit tax rates in their legislations. These rates are set in a simultaneous one-shot game. We use the Nash equilibrium as equilibrium concept. We thus derive the best responses of the two countries.\footnote{The possible alternative would be the Stackelberg case where one country is the first mover. Since it is not clear, however, which country might have this advantage, we consider the more general case of the Nash equilibrium.} So for any possible level of the other countries’ tax rate, we would like to know the best response of the large country and the tax haven, respectively.

**Individual firm behavior and the tax base:** The best response is determined by the way firms react to differences in the tax rates. Whether an individual firm chooses to pay the fixed cost of shifting profits abroad depends on the tax differential and on the level of profits the firm generates. The lower the marginal cost of a firm, the higher are the firm’s profits and thus the more likely it is that the firm chooses to pay the fixed cost of ‘profit-shifting’ FDI.

We define the ‘tax evasion cutoff cost level’ as the cost level \( a^* \) for which a firm is indifferent between paying taxes at home and paying taxes in the tax haven. This cost level is determined by the following condition:

\[
(1 - t_H) \pi(a^*) = (1 - t_X) \pi(a^*) - f_t.
\]

Where \( \pi(a^*) \) are gross profits of a firm with marginal cost of \( a^* \), \( t_H \) is the domestic tax rate, \( t_X \) is the rate set by the tax haven and \( f_t > 0 \) is the fixed costs of becoming a multinational. A necessary condition for this equation to hold is that the difference between the two tax rates \( \rho = t_H - t_X > 0 \). When this condition is violated, no profit shifting takes place: all firms earn higher net profits when paying taxes at home. We first analyze the case of \( \rho > 0 \). Rewriting the tax evasion cutoff condition gives

\[
\pi(a^*) = \frac{f_t}{\rho}
\]

and the ‘tax evasion cost cutoff level’ is

\[
a^* = \left( \frac{\rho \ T_H}{f_t} \right)^{\frac{1}{\pi - 1}}.
\]
The tax base in the home country is given by aggregate profits of firms that have not become multinationals and thus pay taxes at home: \( \Pi_H = \int_{a^*}^{1} \pi(a) \, dG(a) \). The tax base taxed in the tax haven is given by \( \Pi_X = \int_{0}^{a^*} \pi(a) \, dG(a) \). Evaluating the integrals leads to

\[
\Pi_X = \rho \int_T T_2 \tag{9}
\]

\[
\Pi_H = \Pi_H^0 - \Pi_X = T_2 (T_1 - \rho \int_T T_2) \tag{10}
\]

The number of firms paying taxes in the tax haven is given by the measure of firms with marginal cost level below the cutoff level.

\[
N_x = G(a^*) = (a^*)^\gamma \tag{11}
\]

**Proposition 2** Under financial integration when firms have the possibility to do profit shifting FDI, the most productive firms (with a cost level below \( a^* \)) self-select into profit-shifting FDI.

**Proof:** This follows directly from (8), q.e.d

This Proposition states an important result of the model: the most productive firms self-select into profit shifting FDI. In the model with heterogeneous firms, the mass of firms is thus endogenously split into multinationals and domestic firms. Each firm has the possibility to shift profits abroad, but only the most productive firms actually decide to do so. This pattern is in line with the empirical evidence on the determinants of the use of tax haven operations (see Desai, Foley, and Hines (2006)).

### 3.2 Welfare Maximization

**Optimization of the large country** Households have income from labor and receive the net profits of firms paying taxes at home and net profits of firms paying taxes in the tax haven. The aggregate income \( I \) of consumers is given by

\[
I = L + (1 - t_H) \Pi_H + (1 - t_X) \Pi_X - N_x f_t.
\]

Where the last term accounts for the fact that net profits of firms paying taxes in the tax haven are also net of the fixed cost paid to become a multinational.\(^{13}\)

\(^{13}\) We implicitly assume that the fixed cost of tax evasion is not additionally deductible in the tax haven. I.e. the profits to be taxed in the tax haven are not net of the fixed cost.
For any given tax rate of the tax haven \( t_X \), the government of the large country will set its tax rate \( t_H \) to maximize welfare. When it sets \( t_H > t_X \) (i.e., \( \rho > 0 \)), a part of the tax base will flow to the tax haven. When the large country chooses \( t_H \leq t_X \) (i.e., \( \rho \leq 0 \)) there will be no flows of tax base between the two jurisdictions (recall that the tax haven has no tax base of its own).

Define \( U^{\rho>0}(t_H, t_X) \) as welfare in Home for some \( (t_H, t_X) \) combination, conditioned on \( t_H > t_X \) and \( U^{\rho \leq 0}(t_H, t_X) \) as welfare in Home for the opposite case.

The Home country maximizes welfare by choosing for every given \( t_X \) a tax rate \( t_H \) that maximizes \( U(t_H, t_X) = \max \{ U^{\rho>0}(t_H, t_X); U^{\rho \leq 0}(t_H, t_X) \} \). We shall consider \( U^{\rho>0}(t_H, t_X) \) and \( U^{\rho \leq 0}(t_H, t_X) \) separately.\(^{14}\)

**Case 1: \( \rho > 0 \)** We start with the case where for a given \( t_X \), the large country chooses to set a higher tax rate so that \( t_H > t_X > 0 \). Welfare in the large country is then given by

\[
U^{\rho>0} = \bar{U} + (1 - t_H^{\rho>0}) (\Pi_H^A - \Pi_X) + (1 - t_X) \Pi_X - N_X f_t + \beta \ t_H^{\rho>0} (\Pi_H^A - \Pi_X) \tag{12}
\]

where \( t_H^{\rho>0} \) is the tax rate at home conditional on \( \rho > 0 \). The difference \( \Pi_H^A - \Pi_X = \Pi_H \) reflects the fact that some of the tax base in the home country will flow to the tax haven when the tax differential is positive.

From the first order condition (FOC) of the maximization problem, we derive the following implicit response function for a given tax rate of the tax haven

\[
t_H^{\rho>0} = \min \left\{ \frac{(\beta - 1) \left( f_t^\epsilon T_1^{-\epsilon} - \rho f_t^\epsilon \right)}{\epsilon \beta \rho^{-1}}, 1 \right\} \tag{13}
\]

where \( t_X \) and \( t_H \) enter in the tax differential \( \rho \). The following Lemma states a sufficient condition which assures that the second order condition for a maximum is satisfied.\(^{15}\)

**Lemma 1** For a tax rate of the tax haven \( t_X \), (13) is the tax rate of the large country that maximizes (12) if the following sufficient condition holds:

\[
\rho^\epsilon \geq \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1} f_t T_1^{-\epsilon}.
\]

\(^{14}\)We will see later on that neither \( U^{\rho>0}(t_H, t_X) \) nor \( U^{\rho \leq 0}(t_H, t_X) \) dominate for all values of \( t_x \) so that \( U(t_H, t_X) \) is discontinuous. The best response function of the large country will thus also be discontinuous.

\(^{15}\)We limit ourselves to stating a sufficient, not a necessary condition, because we will see later on that this condition is also a necessary and sufficient condition for the large country to set its tax rate according to (13). So the second order condition is satisfied whenever (13) coincides with the best response function of the large country.
Equation (13) provides a relatively simple implicit solution for the optimal tax response function of the government which only depends on the tax differential and parameters of the model. It is of interest to note that the impact of the tax differential crucially depends on $\epsilon$, which only depends on the degree of firm heterogeneity and the elasticity of substitution between varieties.

**Case 2: $\rho \leq 0$** It might be optimal for the large country to set a tax rate smaller or equal to the rate set by the tax haven. In this case all terms related to outflows in the welfare function become zero. Total welfare in this case is:

$$U^{\rho \leq 0} = \bar{U} + \beta t_{H}^{\rho \leq 0} \Pi_{H}^{A} + (1 - t_{H}^{\rho \leq 0}) \Pi_{H}^{A}. \tag{14}$$

For a given level of the tax base it will always be optimal for home to set a tax rate such that it maximizes tax revenue (this is a direct implication of $\beta > 1$). As long as for a given $t_{X}$ the large country sets its tax rate such that $\rho \leq 0$ the tax base will be at its autarky level $\Pi_{H}^{A}$. Conditioned on $\rho \leq 0$, the large country will thus maximize welfare setting

$$t_{H}^{\rho \leq 0} = t_{X} \tag{15}$$

i.e. such that $\rho = 0$.\textsuperscript{16} Next, we will derive the optimal response of the tax haven and then consider the possible equilibria for $\rho > 0$ and $\rho \leq 0$.

**The best response function of the large country:** The objective of the government in the large country is to maximize welfare of its citizens which is given by $U(t_{X}^{H}, t_{H}) = \max \{U^{\rho > 0}(t_{H}, t_{X}^{H}); U^{\rho \leq 0}(t_{H}, t_{X}^{H})\}$. We have seen above that (provided the condition in Lemma 1 holds) $t_{H}^{\rho > 0}$ given by (13) maximizes $U^{\rho > 0}(t_{H}, t_{X}^{H})$ and $t_{H}^{\rho \leq 0}$ from (15) maximizes $U^{\rho \leq 0}(t_{H}, t_{X}^{H})$.

The large country will set $t_{H}$ according to (13) as long as $U(t_{H}^{\rho > 0}, t_{X}) \geq U(t_{H}^{\rho \leq 0}, t_{X})$. The following Lemma states the condition on the tax differential for this to be the case.

\textsuperscript{16}Note that in the autarky equilibrium, the large country sets the highest possible tax rate under the condition that $t_{A}^{H} \leq 1$. Here the maximization is conditioned on $\rho \leq 0$ so that that highest possible tax rate the large country can set is just the tax rate set by the tax haven.
Proposition 3 Under financial integration when firms have the possibility to do profit shifting FDI,

(i) the best response of the large country for a given \( t_X \) is given by (13) as long as the implied value of \( t_H \) is large enough to satisfy

\[
\rho^e \geq \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1} f_H^e T_1^{-\epsilon}. \tag{16}
\]

(ii) If for a given \( t_X \) the value of \( t_H \) implied by (13) is too small to satisfy this condition, the best response of the large country is given by (15), which implies \( \rho = 0 \).

Proof: see appendix B.

Note that condition (16) is the same as the one used in Lemma 1. This implies that when equation (13) coincides with the best response function (and only then it is relevant for the analysis), the SOCs for a maximum hold.

The following Corollary to Proposition 3 states for which values of the tax haven’s tax rate the best response of the large country is given by (13) and for which values it is given by (15).

Corollary 1 Under financial integration when firms have the possibility to do profit shifting FDI, (13) is the best response of the large country to all values of the tax rate of the tax haven that satisfy

\[
t_X \leq \frac{(\beta - 1) f_H^e T_1^{-\epsilon}}{\epsilon \beta} - \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1 + (\beta - 1)} f_H^e T_1^{-\epsilon} \left( \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1 + (\beta - 1)} f_H^e T_1^{-\epsilon} \right)^{1/\epsilon}. \tag{17}
\]

For all values of \( t_X \) that do not satisfy this condition, (15) is the best response of the large country.

Proof: see appendix C.

While conditions (16) and (17) might look a bit complicated, it is important to note that they only depend on preference parameters, the fixed cost of becoming a multinational and the degree of firm heterogeneity.

Optimization of the tax haven: As outlined above, the structure of the tax haven is kept as simple as possible. It is a jurisdiction outside the large country that has no own tax base. It
maximizes tax revenue from multinational companies with an affiliate in the tax haven. Taking the tax rate in the large country as given, the tax haven maximizes $V = t_X \Pi_X$. In principle the tax haven could set its tax rate below or above the tax rate of the big country. When it sets a tax rate equal or above the tax rate of the large country ($\rho \leq 0$), it will have zero tax revenue. Thus, for any positive tax rate of the large country, it will always be optimal for the tax haven to undercut the large country, so that $\rho > 0$. Revenue maximization leads to

$$t_X = \min \left\{ \frac{\rho}{\epsilon} , \frac{1}{\epsilon+1} \right\} = \frac{t_H}{\epsilon+1} = \frac{\gamma}{\sigma-1} t_H.$$  \hspace{1cm} (18)

Note that $\epsilon+1 = \gamma/(\sigma-1) > 1$. The tax haven thus sets a tax rate that is a constant fraction of the tax rate of the large country. The extent to which the tax haven undercut the large country is determined by the degree of firm heterogeneity and the elasticity of substitution. We can now state the following proposition:

**Proposition 4** Under financial integration when firms have the possibility to do profit shifting FDI,

(i) the best response of the tax haven for a given $t_H > 0$ is given by (18) and implies that the tax haven will always undercut the large country.

(ii) the undercutting is the stronger the lower the firm heterogeneity (the higher $\gamma$) and the stronger the market power of individual firms (lower $\sigma$).

(iii) the best response of the tax haven for $t_H = 0$ is indeterminate.

**Proof:** (i) follows from the fact that for $t_H > 0$, a tax rate of $t_X \geq t_H$ implies $\Pi_X = 0$ and thus $V = 0$, while any $0 < t_X < t_H$ implies $\Pi_X > 0$ and thus $V > 0$. (ii) follows directly from (18). For $t_H = 0$ any $t_X \in [0, 1]$ implies $\Pi_X = 0$ and thus $V = 0$, which proves (iii). q.e.d.

### 4 Equilibrium

In this section the equilibrium of the model is derived and the effect of firm heterogeneity and market structure on existence and shape of the equilibrium is discussed.

#### 4.1 Equilibrium of the Tax Game

Propositions 3 and 4 characterize the best response functions of the two governments. Any intersection of the two best response functions is a Nash equilibrium. In principle, there could
be equilibria for $\rho > 0$ and for $\rho \leq 0$. Proposition 4 implies, however, that for any $t_H > 0$ the tax haven set its tax rate as a fraction of the tax rate of the large country. So any $(t_H; t_X)$ combination which implies $\rho \leq 0$ cannot be an equilibrium.

If an equilibrium exists, it has to be in the range where (13) describes the best response of the large country i.e. where $\rho > 0$. The only possible equilibrium is thus at the intersection of (13) and (18).

Assume that a unique intersection exists. Label $t^e_X$ and $t^e_H$ the tax rates at the intersection. They will be such that $t^e_X$ maximizes $V(t_X, t_H^e)$ for a given $t^e_H$, and given $t^e_X$, the large country maximizes $U^{\rho > 0}(t^e_X, t_H)$ by choosing $t^e_H$. This intersection, however, will only be a Nash equilibrium if the best response to $t^e_X$ is indeed given by (13) and not by (15). This is the case when $t^e_X$ fulfills condition (17).\(^\text{17}\)

The following Lemma summarizes this result.

**Lemma 2** If an intersection of (13) and (18) exists, it is a Nash Equilibrium of the tax game if and only if for the tax rate of the tax haven at the intersection ($t^e_X$), the $t^e_H$ implied by (13) is large enough to satisfy (16). Equivalently, if $t^e_X$ satisfies condition (17).

**Proof:** This follows directly from Proposition 3 (i) and (ii), Corollary 1 and the definition of a Nash equilibrium. q.e.d.

In the following analysis we will assume that this parameter condition holds.

**The equilibrium:** We will now determine the the equilibrium of the tax game. Taking the difference between (13) and (18) we can solve for the equilibrium tax differential:

$$\rho^* = \left(\frac{\beta - 1}{\epsilon \beta + 2 \beta - 1}\right) \frac{f_t}{T_1}. \quad (19)$$

It directly follows that the optimal tax rates for $\rho > 0$ are given by

$$t^*_X = \frac{1}{\epsilon} \left(\frac{\beta - 1}{\epsilon \beta + 2 \beta - 1}\right)^\frac{1}{2} \frac{f_t}{T_1} \quad \text{and} \quad t^*_H = \frac{\epsilon + 1}{\epsilon} \left(\frac{\beta - 1}{\epsilon \beta + 2 \beta - 1}\right)^\frac{1}{2} \frac{f_t}{T_1} \quad (20)$$

The level of the two equilibrium tax rates are determined by parameters of the model only. Namely preference parameters, the fixed cost of FDI and the degree of firm heterogeneity.

Based on these results, we can derive the equilibrium cutoff productivity in equilibrium:

$$a^{**} = \left(\frac{\beta - 1}{\epsilon \beta + 2 \beta - 1}\right)^{\frac{1}{\sigma - 1}}. \quad (21)$$

\(^{17}\)Graphical intuition is provided in the numerical analysis in section 5.
The equilibrium number of firms choosing ‘profit shifting’ FDI is

\[ N_X^* = \left( \frac{\beta - 1}{\epsilon \beta + 2 \beta - 1} \right)^{\frac{\epsilon + 1}{\alpha}}. \]  \tag{22}

The part of the tax base that flows to the tax haven in equilibrium is

\[ \Pi_X^* = \frac{\beta - 1}{\epsilon \beta + 2 \beta - 1} \frac{\alpha}{\sigma}. \]  \tag{23}

To give some intuition, the following subsection provides some graphical examples of best response functions calculated for specific parameter values. They illustrate the equilibrium for different parameter constellations (in particular \( \epsilon > 1 \) and \( \epsilon < 1 \)) as well as the case when no equilibrium exists.

\section*{4.2 Heterogeneity and Existence of the Equilibrium}

Condition (16) determines whether the best response of the large country is given by equation (13). For \( \epsilon > 1 \) this condition is always met. When \( \epsilon < 1 \) it is possible that no equilibrium exists. Figures 1, 2 and 3 in the appendix provide some graphical intuition to this point.

Figure 1 illustrates a case of \( \epsilon > 1 \). For all possible tax rates of the tax haven the best response of the large country is given by (13) (solid line). As stated in Proposition 4 for a given positive \( t_H \) the tax haven will always undercut the large country. This is reflected by its best response function (dashed line). From equation (18) the slope of this best response function is inversely related to \( \epsilon \).

Figure 2 represents the case where \( \epsilon < 1 \) and the equilibrium exists. The graph shows that the response function of the large country is discontinuous. For low values of \( t_X \) (for all values that satisfy (17)) it is optimal for the large country to set a higher tax rate according to (13). For higher tax rates of the tax haven it is optimal for the large country to set the same tax rate as the tax haven.

This discontinuity in the best response function is an important feature of the model. As derived in the theoretical analysis, for a given \( t_X \) the best response of the large country falls in one of two categories. It can set its rate higher than the tax haven (according to (13)) or it sets the same rate as the tax haven. When it sets a higher rate it has higher revenues for each unit of tax base remaining in the country (intensive margin) but it faces an outflow of tax base to the tax haven (extensive margin). For high values of \( t_X \) the large country can set a relatively high tax rate. In this case the intensive margin is less important and the extensive margin dominates.
This causes the large country to set its rate equal to the rate of the tax haven to avoid outflows of the tax base.

If and when this discontinuity in the response function occurs depends on the relative importance of the intensive and extensive margin. The more elastic the tax base with respect to differences in the tax rate the more important becomes the extensive margin. This elasticity and its determinants will be discussed in detail below. As will be shown in section 5, when \( \epsilon \) is low the elasticity of the tax base is high. So the lower \( \epsilon \), the lower will be the tax rate \( t_X \) (given by (17)) for which the large country switches to the \( \rho = 0 \) part of its best response function.

Figure 3 illustrates this point graphically. A lower \( \epsilon \) (higher elasticity) leads the large country to switch the the \( \rho = 0 \) part of its best response function already for low values of \( t_X \). Since the tax haven always undercuts (it never chooses \( \rho = 0 \)) it is possible that no equilibrium exists. This is the case illustrated in Figure 3.

5 The Role of Firm Heterogeneity and Market Power

As discussed before all firms with a unit cost below the cutoff cost level in equation (8) opt for ‘profit shifting’ FDI when a positive tax difference emerges. For the tax game the distribution of profits across different cost levels is thus crucial. In this context it is especially important whether the most productive firms (the first firms to shift profits) account for a large fraction of aggregate profits. In this section we show how the distribution of firm profits is determined by the degree of firm heterogeneity and the extent of market power in the economy. Furthermore we show that the elasticity of the tax base also depends on these two determinants. Finally we analyze the impact of firm heterogeneity on the equilibrium outflow of tax base and on equilibrium tax rates.

5.1 Heterogeneity, Market Power and the Tax Base

In this subsection we will investigate the role of firm heterogeneity and market power on the the distribution of aggregate profits (the tax base) across firms. Define \( \Pi(a) \) as the share of aggregate profits accounted for by firms with productivity level \( a \): \( \Pi(a) \equiv \frac{\pi(a)g(a)}{\Pi g} \). It can be shown that \( \frac{\partial \Pi(a)}{\partial a} = (\gamma - \sigma)(\gamma - (\sigma - 1))a^{\gamma - \sigma - 1} \). The sign of this partial derivative (the slope of \( \Pi(a) \) in \( a \)) is positive if \( \gamma > \sigma \) (a case considered in Figure 4). It is zero or negative if \( \gamma \leq \sigma \) (a case considered in Figure 5).

Figure 4 and 5 provide a graphical illustration of the distribution of the tax base across firms. The solid line in the graphs represents aggregate profits generated by firms with a cost level of \( a \).
One point on the curve represents the share of overall profits that is accounted for by firms with some cost level $a$. The thick dashed line plots the measure of firms with the different productivity levels and the thin vertical dashed line marks the equilibrium cutoff productivity level given by (21). The area under the solid curve to the left of the $a^{**}$ line represents the fraction of aggregate profits shifted to the tax haven $\Pi_X^H$. While the area under the dashed line represents the measure of firms that become multinationals.

Figure 4 illustrates the case where $\gamma > \sigma$. In the first graph we have $\sigma = 5$ and $\gamma = 6$. Aggregate profits by cost level ($\Pi(a)$) increase in the cost level, which implies that the smaller and less productive firms contribute more to the tax base than the large and very productive ones. This is the case as for $\gamma > \sigma$ the measure of firms grows faster than the per firm profits decline. Comparing the dashed and the solid line to the left of $a^{**}$ shows that a very small measure of very productive firms accounts for the considerable fraction of the tax base which is transferred to the tax haven.

The comparison with the other graphs illustrates how the degree of firm heterogeneity $\gamma$ and the elasticity of substitution $\sigma$ affect the tax base and the equilibrium cutoff. In the graph on the right $\sigma$ remains unchanged but $\gamma$ increases to 8. The higher $\gamma$ (lower firm heterogeneity: most of the firms have similar high cost levels) increases the measure of firms with low productivity. Their increased number explains the increase in the fraction of the tax base the low productivity firms account for. Higher productivity firms account for less of the aggregate profits and the cutoff level increases.

The third graph shows that a decrease in $\sigma$ (an increase in the monopolistic market power of individual firms) has a similar effect. Keeping $\gamma$ constant and reducing $\sigma$ to 3.5, the low productivity firms account for an even higher share of aggregate profits. The intuition is that since consumers are less able to substitute the goods of the high cost producers, the profits of this group of firms (relative to aggregate profits) rise. Again, the equilibrium cutoff level decreases.

The first graph in Figure 5 illustrates the case where $\sigma = \gamma = 6$. In this special case aggregate profits are uniformly distributed across cost levels. Here the higher profits made by very productive firms are exactly offset by their lower measure. The second graph shows the case where the economy is very favorable to very productive firms. Although there are only few of them, their profits are so high that they represent an important fraction of aggregate profits.

---

18 We have seen above that the level of aggregate profits $\Pi^H$ depends on the elasticity of substitution $\sigma$. We normalize profits by $\Pi^H$ so that the area under the solid curve is unity in all cases even when the $\sigma$ differs across graphs.
5.2 The Elasticity of the Tax Base

As shown above industry structure matters for government policies. To gain a better understanding of the role industry- and market structure play in the decision process of the two governments it is of interest to look at the elasticities of the tax base the governments take into account when setting their tax rates.

For a given tax rate of the large country the tax haven faces an elasticity of its tax base $\Pi_X$ of

$$
\eta^\pi = - \frac{\partial \Pi_X}{\partial t_X} \frac{t_X}{\Pi_X} = \epsilon \frac{t_X}{t_H - t_X} = \left( \frac{\gamma}{\sigma - 1} - 1 \right) \frac{t_X}{t_H - t_X}.
$$

(24)

It is thus determined by three factors: the degree of firm heterogeneity, the substitutability between goods and the tax rate of the tax haven.

For a $t_X$ close to $t_H$, the elasticity becomes very large and increases in $t_X$. Revenue maximization induces the tax haven to decrease its tax rate until the elasticity equals unity. At this point the (negative) intensive margin effect (tax revenues per unit of tax base taxed) and the (positive) extensive margin effect (revenue from taxing a larger tax base) exactly offset each other.

The higher $\epsilon$, the lower the level of $t_X$ necessary to equate the elasticity to unity. A low heterogeneity (high $\gamma$) thus implies ceteris paribus a high value of the elasticity and thus leads the tax haven to set a low tax rate. The same holds true for strong monopolistic market power (low $\sigma$). Thus, the undercutting of the tax haven is the stronger the lower firm heterogeneity and the stronger market power.

This corresponds to the third graph in Figure 4. The productive firms only account for a small part of the tax base so the tax haven has to set a low tax rate in order to attract some of the profits generated by firms with intermediate costs levels. In equilibrium this policy results in a higher cutoff cost level. The opposite is the case in the second graph in Figure 5 where firm heterogeneity is strong and market power low. In this case even a moderate degree of undercutting allows the tax haven to attract a large fraction of the tax base as high productivity firms account for a large fraction of profits. As a result the tax haven undercuts less, which implies a low cutoff cost level.

5.3 Heterogeneity, Equilibrium Profit Shifting and Tax Rates

A crucial question for understanding actual tax policies in different countries is how firm heterogeneity affects the optimal policy of the large country. Put differently: are countries with some type of industry structure more affected by the pressures of international tax competition than
The model implies that this is clearly the case: the more homogeneous firms (high $\gamma$) the less of the tax base is induced to leave the country. Figure 6 illustrates this important finding. The solid line represents the fraction of the tax base flowing to the tax haven. These outflows are unambiguously decreasing in $\gamma$. In line with the intuition above, when the most productive firms account for a small fraction of aggregate profits only, equilibrium outflows are lower.

In line with the above intuition, Figure 7 illustrates that the difference between the two equilibrium tax rates is increasing in $\gamma$. The higher $\gamma$ (low heterogeneity) the stronger the tax haven undercuts and the less the large country deviates from its autarky tax rate.

These results imply that countries with more homogeneous industry structures are ceteris paribus less affected by international tax competition. A possible example is Germany which has a large number of small and medium size enterprises (the German ‘Mittelstand’) that produce highly specialized products. Our model suggests that Germany should be less worried about international tax competition than other countries in which a few large firms account for a large fraction of aggregate profits.

6 Behind the Scenes: Our Assumptions

We have chosen a very particular set of assumptions to construct this model. These assumptions are not at all randomly chosen, but their very combination allows us to make progress in an area where the literature has so far failed to deliver reliable results. To the best of our knowledge, very few attempts have been made to combine heterogeneous firms and monopolistic competition in order to determine their role for setting taxes in an environment of international tax competition. To our knowledge, our paper is the only one to derive an equilibrium in such a setting staying in line with the standard assumptions of economic modeling. (TWO REFERENCES HERE).

We consider it one contribution of the paper to outline a way standard assumptions can be combined in order to avoid the numerous caveats in the modeling process. We will thus give a brief discussion of the specific assumptions we have made and outline why they are necessary to achieve analytical tractability.

Models with firm heterogeneity and monopolistic competition à la Melitz (2003) have interesting implications for firm entry and exit. Before becoming a firm, a pool of potential entrants

\footnote{We have seen that market structure is the second crucial element. When a country has specialized in industries with low substitutability it should also be less affected by international tax competition.}
computes discounted expected (net) profits of starting a firm and there will be entry until this is equalized to the fixed cost of entry. This leads to an endogenous number of firms in an industry. The problem with introducing a profit tax is that this tax influences net profits and thus firm entry. When setting the optimal tax rate the government would thus have to take into account the effect of the tax rate on firm entry. In addition, and from an analytical point more problematic, the tax rate also affects the cutoff cost level i.e. the level of marginal cost below which it is profitable for a firm to produce (in order to recover the fixed cost of production). The cutoff level influences almost all important variables in the system via the aggregate price index, which depends on the tax rate in a highly non-linear way. Although the system of equations and unknowns is well defined, there is little hope to determine best response functions of the governments that can be used in the tax game.

In order to deal with these problems, we make the following simplifying assumptions. Similar to Chaney (2008), we assume a fixed and exogenous number of firms in a country. In addition we assume a fixed cost of operation of zero. Taken together, these assumptions assure that the tax rate does not influence the number of active firms, leaves their price setting and thus the price index unaffected.

The problem that arises in such a setting is that net profits need to be redistributed to the owners of the firms. But since net profits depend on the tax rate, the income of the households also depends on the tax rate. Under the standard Cobb-Douglas-CES preferences expenditure on the differentiated goods is not constant, but a constant fraction of income. Expenditure on the differentiated goods is thus depending on the tax rate. Via the expenditure, the tax rate affects all equilibrium variables again in a very non-linear way, so that only implicit solutions can be found that cannot be used in the tax game.

In order to avoid this problem we use quasi-linear preferences like Melitz and Ottaviano (2008) and Baldwin and Okubo (2006). Under appropriate assumptions, these preferences assure that the expenditure on the differentiated good sector is constant even when income fluctuates.

To summarize, the crucial assumptions in order to get closed form solutions for the closed-economy in our model are: a fixed number of firms, zero fixed costs and quasi-linear preferences. All these assumptions have been used in the literature before and none of these assumptions appears to invalidate our results on the role of firm heterogeneity and market structure.

In standard NEG models of tax competition, tax differences lead to relocation of firms, under costly trade (or: zero trade) this represents an additional channel how the tax rates affect the
price index and thus the whole system of equilibrium conditions. We do not have this problem in our model, because we consider profit-shifting FDI only. So even if a firm decides to pay taxes abroad, it will still charge the same prices and will still produce in the home market. So the price indices remain unaffected.

Finally, the modeling of the public good in the utility function is crucial. The way the public good enters the consumer’s utility has to be such that there is some reason for the government to spend resources on it. If the public good were strictly dominated by the homogeneous good, the government would just leave the profits untaxed to its consumers. It certainly is a strong assumption not to introduce any convexity of the utility generated by the government into the utility function. Under the $\beta > 1$ assumption, the government will always try to spend as many resources as possible on the public good. If it can, it will thus always set a tax rate of 100%. Although this implication is not very realistic, this is the most parsimonious way to have a desirability of a public good combined with the possibility of a race to the bottom. Numerical robustness checks show that using $\ln G$ instead of $\beta G$, does not qualitatively change the results.

7 Conclusions

In this paper we have proposed a stylized model of international tax between a large country and a tax haven. The methodological contribution is to provide a fully solvable model of international tax competition with heterogeneous firms and monopolistic competition.

Our analysis reveals that firm heterogeneity and monopolistic market power are key for the tradeoffs the governments are facing. While all firms have the possibility to do ‘profit shifting’ FDI, only the most productive firms do so. It is then crucial for the governments whether these firms account for a large fraction of the tax base (aggregate profits). This is the case when firms are very heterogeneous (relatively many relatively productive firms) and monopolistic market power is low (high degree of substitutability between goods). A country with such an industry structure will be strongly affected by international tax competition.

The opposite holds true for economies with a low degree of firm heterogeneity and a high degree of monopolistic market power. These economies are ‘shielded’ from international tax competition. In equilibrium they face lower outflows of the tax base and can set higher tax rates.

The tractability of the model, the clear predictions on individual firm behavior and the role of firm heterogeneity in tax competition open some interesting fields for future research. On the theoretical side ongoing work assesses the potential to extend the model to a multi (large)
country setting.
The model predicts that the most productive firms choose ‘profit shifting’ FDI. This could be investigated empirically. Also on the empirical side, an analysis of the relation of industry structure and the reaction to increasing international tax competition are an interesting field of research.

The model could also provide a foundation for policy analysis. For example it implies that tax rate harmonization (e.g. in the European Union) has asymmetric effects on equilibrium tax revenues when the industry structures in the participating countries differ. According to the model, countries like Germany with a large number of small and medium size enterprises producing highly specialized goods should be relatively ‘shielded’ from international tax competition, allowing to set relatively high tax rates. But countries in which very productive firms account for a large fraction of aggregate profits optimally set lower tax rates. Tax harmonization would affect these countries differently.
References


Sato, Yasuhiro, and Jacques-François Thisse, 2007, “Competing for Capital when Labor is Het-

Slemrod, Joel, and John D. Wilson, 2006, “Tax Competition With Parasitic Tax Havens”, NBER 

Economics* 19, 296–315.


Zodrow, G.R., and P. Mieszkowski, 1986, “Pigou, Tiebout, Property Taxation, and the Under-
Appendix

A Proof of Lemma 1

First note that the second derivative of the welfare function with respect to the tax rate of the large country is given by

$$\frac{\partial^2 U}{\partial t_H^2} = -2(\beta - 1) \frac{\partial \Pi_X}{\partial t_H} - \frac{\partial^2 \Pi_X}{\partial^2 t_H} [t_H + t_X] - \frac{\partial^2 N_x}{\partial^2 t_H} f_t.$$  

For notational convenience, we define $T_5 \equiv T_2 e^{j-1} f_t^{-1} > 0$. We then use $\frac{\partial \Pi_X}{\partial t_H} = T_5$ together with $\frac{\partial^2 \Pi_X}{\partial^2 t_H} = \frac{1}{\rho} T_5$ and $\frac{\partial^2 N_x}{\partial^2 t_H} f_t = \epsilon T_5$ to get

$$\frac{\partial^2 U}{\partial t_H^2} = -T_5 \left( 2(\beta - 1) + \epsilon + \frac{1}{\rho} ((\beta - 1)t_H + t_X) \right) = -T_5 \left( 2(\beta - 1) + \epsilon - (\epsilon - 1) + (\epsilon - 1) \frac{t_H}{\rho} \right).$$

From (13) it follows that $T_5 = \frac{[\beta - 1] f_t^{-1} T_1^{-\epsilon}}{\epsilon \beta \rho^\epsilon} - \frac{[\beta - 1]}{\epsilon \beta}$ so that

$$\frac{\partial^2 U}{\partial t_H^2} = -T_5 \left( 2(\beta - 1) + \epsilon + \frac{(\beta - 1)(\epsilon - 1)}{\epsilon} \left( \frac{f_t T_1^{-\epsilon}}{\rho^\epsilon} - 1 \right) - \epsilon + 1 \right). \quad (25)$$

To prove the Lemma, we need to show that

$$\rho^\epsilon \geq \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1} \frac{f_t T_1^{-\epsilon}}{\rho^\epsilon}.$$  

is a sufficient condition for (25) to be negative. To do so, we proceed in two steps. We first show that it is negative when the above condition holds with equality. We then show that this also holds true for larger values of $\rho$.

Define $\rho^*$ as the value of $\rho$, where the above condition holds with equality. We will first determine the sign of (25) for $\rho^*$ i.e. $\frac{\partial^2 U}{\partial t_H^2} |_{\rho^*}$.

$$\frac{\partial^2 U}{\partial t_H^2} |_{\rho^*} = -T_5 \left( 2(\beta - 1) + 1 - T_j - \frac{(\beta - 1)(\epsilon - 1)}{\epsilon} \right) \quad (26)$$

With $T_j = \frac{(\beta - 1)(1 - \epsilon)}{\epsilon} \frac{\frac{\epsilon}{\epsilon + 1} + (\beta - 1)}{(\beta - 1)(1 - \epsilon)} = \frac{1}{\epsilon + 1} + \frac{\beta - 1}{\epsilon} > 0$
Simplifying and recalling that $T_5 > 0$ then gives:

$$ \frac{\partial^2 U}{\partial t_H^2} |_{\rho^*} = -T_5 \left( (\beta - 1) + \frac{\epsilon}{\epsilon + 1} \right) < 0. \tag{27} $$

This implies that for $\rho = \rho^*$ the second order condition holds and (13) is indeed the optimal response.

To see that this is generally true as long as $\rho \geq \rho^*$, note that any value of $\rho \geq \rho^*$ can be written as $\rho = x \rho^*$ with $x \geq 1$. In order to obtain $\frac{\partial^2 U}{\partial t_H^2} |_{\rho^*}$, we have plugged in $\rho^*$ into (25). Now considering any value of $\rho \geq \rho^*$, we can plug $\rho = x \rho^*$ into to (25).

It can be seen in (25) that $\rho$ enters twice in the expression for $\frac{\partial^2 U}{\partial t_H^2}$. Entering via $T_j$ it does not affect the sign. To see the effect of a higher $\rho$ on the second term, note when we use $\rho = x \rho^*$, instead of $T_j$ in equation (26), we would now have $T_j \frac{1}{\epsilon} < T_j$. The positive effect of $T_j$ on the sign of $\frac{\partial^2 U}{\partial t_H^2}$ is thus dampened for any $\rho > \rho^*$. This shows that the condition stated in the Lemma is indeed a sufficient condition for (13) to be a utility maximum. q.e.d.

## B Proof of Proposition 3

To see (i), note that the large country will set $t_H$ such that $\rho > 0$ (i.e. according to (13)) as long as $U(t_H^{\rho > 0}, t_X) \geq U(t_H^{\rho < 0}, t_X)$. Plugging in (12) and (14) into this condition, rearranging and using (15) and $(1 - t_X) - (1 - t_H^{\rho > 0}) = \rho$ we get:

$$ \rho \left( \beta - 1 \right) \Pi_H^A t_H^{\rho > 0} - \rho \right) \Pi_X + N_X \ f_t. \tag{28} $$

Now first using $f_t N_x = \Pi_X \rho \frac{\epsilon}{\epsilon + 1}$, then $\Pi_H^A = T_2 \ T_1^{-\epsilon}$ and $\Pi_X = \rho^c \ T_2 \ f_t$, simplifying and solving for $t_H^{\rho > 0}$ gives:

$$ t_H^{\rho > 0} \leq \frac{(\beta - 1)T_1^{-\epsilon} f_t + \rho^c \frac{1}{\epsilon + 1}}{\beta \ \rho^{\epsilon - 1}}. \tag{29} $$

To see for which values of $\rho$ the tax rate in (13) satisfies condition (29), we plug (13) into (29), which gives:

$$ \rho^c \geq \frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1} \ f_t T_1^{-\epsilon}. $$

As long as for a given $t_X$ the $t_H$ implied by (13) is high enough to satisfy this condition, we will have $U(t_H^{\rho > 0}, t_X) \geq U(t_H^{\rho < 0}, t_X)$ and thus the best response function of the large country given
by (13). This proves (i).

(ii) directly follows from the fact that when the above condition is violated, we have $U(t_H^{\rho > 0}, t_X) < U(t_H^{\rho < 0}, t_X)$. In this case the best response of the large country is given by (15) because it maximizes $U(t_H^{\rho < 0}, t_X)$ and implies $\rho = 0$. q.e.d.

C Proof of Corollary 1

Define $\rho^{\text{jump}}$ as the tax differential just before the regime switch to $\rho = 0$. For $\rho^{\text{jump}}$, (16) holds with equality. Note that $t_X^{\text{jump}} = t_H^{\rho > 0}[\rho^{\text{jump}}] - \rho^{\text{jump}}$ combining this with (13) and (16), the value of $t_X$ for which the best response for the large country switches from $\rho > 0$ to $\rho = 0$ is given by

$$t_X^{\text{jump}} = \frac{(\beta - 1)f_T^{-\epsilon}}{\epsilon\beta \left( \frac{1-\epsilon}{\epsilon+1} \right)^{1/\epsilon}} - (\beta - 1) + \epsilon\beta \left( \frac{1-\epsilon}{\epsilon+1} + (\beta - 1) f_T^{-\epsilon} \right)^{1/\epsilon}.$$

To prove the inequality in Corollary 1, it remains to be shown that all values of $t_X$ below $t_X^{\text{jump}}$ imply that the large country sets its tax rate such that $\rho > 0$.

We know that the utility of the best response with $\rho > 0$ dominates for $t_X = t_X^{\text{jump}}$. A sufficient condition for this to hold for $t_X \leq t_X^{\text{jump}}$ as well, is that $\rho$ decreases in $t_X \forall t_X \leq t_X^{\text{jump}}$. From before we have:

$$t_H = \frac{(\beta - 1)(f_T^{-\epsilon} - \rho)}{\epsilon\beta \rho^{\epsilon-1}}$$

subtracting $t_X$ on both sides and multiplying by $\epsilon\beta \rho^{\epsilon-1}$ we get:

$$\epsilon\beta \rho^{\epsilon} = (\beta - 1)f_T^{-\epsilon} - (\beta - 1)\rho - t_X \epsilon\beta \rho^{\epsilon-1}$$

This can be rewritten as:

$$t_X = Q(\rho) = \frac{1}{\epsilon\beta} \left[ (\beta - 1)f_T^{-\epsilon} \right] \rho^{1-\epsilon} - (\epsilon\beta + \beta - 1)\rho$$

It remains to show that $Q'(\rho) < 0 \forall t_X \leq t_X^{\text{jump}}$

$$Q'(\rho) = (1 - \epsilon)(\beta - 1)f_T^{-\epsilon} \rho^{1-\epsilon} - (\epsilon\beta + \beta - 1) < 0$$

This condition can be rewritten to:

$$\rho^{\epsilon} > \frac{(1 - \epsilon)(\beta - 1)f_T^{-\epsilon}}{(\epsilon\beta + \beta - 1)}$$
Condition (6) gives a lower bound for $\rho$. Plugging in this bound into the previous condition delivers:

\[
\frac{(1 - \epsilon)(\beta - 1)}{\epsilon + 1 + \beta - 1} f_{\epsilon T_1^{-\epsilon}} > \frac{(1 - \epsilon)(\beta - 1)}{(\epsilon \beta + \beta - 1)} f_{\epsilon T_1^{-\epsilon}}
\]

which can be simplified to:

\[
\beta > \frac{1}{\epsilon + 1}
\]

which is always true. q.e.d.
D Figures

Figure 1: This Figure provides a numerical example for the equilibrium of the tax game for $\epsilon > 1$. In this case the best response function of the large country is continuous (bold solid line) and there exists a unique intersection with the best response function of the tax haven (dashed line). The best response of the big country implies $\rho > 0$ for all tax rates of the tax haven. The parameter values chosen are $\sigma = 4$, $\gamma = 6.6$ (implying $\epsilon = 1.2$) and $f_t = 0.2$. 
Figure 2: This numerical example illustrates that for $\epsilon < 1$ the reaction function of the large country (bold solid line) is discontinuous. A low epsilon implies a strong elasticity of the tax base. For high values of $t_X$ the best response of the large country is to set the same tax rate as the tax haven, which implies a zero outflow of tax base. In this example with $\epsilon = 0.9$ the equilibrium exists. The parameter values chosen are $\sigma = 4$, $\gamma = 5.7$ (implying $\epsilon = 0.9$) and $f_t = 0.5$.

Figure 3: This numerical example uses the same parameter values as Figure 2 except that $\gamma$ (and thus $\epsilon$) is lower. I.e. firm heterogeneity is stronger implying a higher elasticity of the tax base. Even for low levels of $t_X$ the large country sets the same tax rate as the tax haven. In this case no equilibrium exists. The parameter values chosen are $\sigma = 4$, $\gamma = 4.8$ (implying $\epsilon = 0.6$) and $f_t = 0.5$. 
Figure 4: These three graphs illustrate the effect of firm heterogeneity and market power on the distribution of the tax base (solid line). Profits are normalized by overall aggregate profits (\( \Pi_H^A \)) thus the area under the curve is unity. A point on the curve represents the share of overall profits firms with cost level \( a \) account for. The dashed line plots the density of firms \( g(a) \). In all three graphs \( \gamma > \sigma \). In the first graph \( \sigma = 5 \) and \( \gamma = 6 \). In the graph on the right \( \gamma \) is shifted to 8. In this case there is a larger mass of firms with high cost levels (dashed line). These firms now account for a larger share of profits. In the third graph, in addition to the shift in gamma, \( \sigma \) is lowered to 3.5 which further increases the share of the high cost firms in overall profits.
Figure 5: These graphs illustrate the cases of $\sigma = \gamma$ and a case where $\sigma > \gamma$. In the first graph the effect of the lower number of high productivity firms is exactly offset by their higher profits. This implies that aggregate profits are distributed uniformly across cost levels. In the second graph $\sigma = 6$ and $\gamma = 5.7$. In this case there are relatively many very productive firms (low $\gamma$) which earn relatively high profits (high $\sigma$). Thus a large fraction of aggregate profits is accounted for by very productive firms.

Figure 6: This Figure provides an illustration of the role of firm heterogeneity for profit shifting, i.e. for the fraction of the tax base flowing to the tax haven (solid line) and the fraction of firms that choose ‘profit shifting’ FDI (dashed line). An increase in $\gamma$ implies that firms are more homogeneous and thus that the elasticity of the tax base decreases. The lower this elasticity becomes (the higher $\gamma$ and thus the higher $\epsilon$) the lower is the fraction of the tax base the tax haven can attract. Note that the graph is only valid for values of $\gamma$ of about 6 and higher. For lower values the existence condition of the equilibrium is violated.
The effect of heterogeneity on equilibrium tax rates

Figure 7: This graph shows how the two equilibrium tax rates are affected by the degree of firm heterogeneity. The higher $\gamma$ the lower is the fraction of of highly productive firms and thus the lower is the elasticity of the tax base. The graph shows that for higher $\gamma$ the large country sets a higher tax rate. Furthermore the tax difference rho increases in gamma, while the tax rate of the tax haven first increases and then decreases in gamma. This is an important result of the model: The more homogeneous the industry structure the less the large country is affected by the presence of the tax haven. Countries with a more homogenous industry structure can set their tax rate closer to its first best. Overall the negative effects of a tax haven on welfare in the big country are smaller. Note that the thin dashed line represents the cutoff $t_X$ in equation (17). The graph is only valid to the right of the intersection of the two dashed lines (before no equilibrium exists).